Available online at www.sciencedirect.com

ScienceDirect

Karbala International Journal of Modern Science 2 (2016) 156–168

http://www.journals.elsevier.com/karbala-international-journal-of-modern-science/

DNA barcoding and evolutionary lineage of 15 insect pests of horticultural crops in South India

Pushparaj Karthika [a,*], Natraj Krishnaveni [a], Chithravel Vadivalagan [b], Kadarkarai Murugan [b], Marcello Nicoletti [c], Giovanni Benelli [d]

[a] *Department of Zoology, Faculty of Science, Avinashilingam Institute for Home Science and Higher Education for Women, 641043, Tamil Nadu, India*

[b] *Division of Entomology, Department of Zoology, School of Life Sciences, Bharathiar University, Coimbatore, 641 046, Tamil Nadu, India*

[c] *Department of Environmental Biology, Sapienza University of Rome, Piazzale Aldo Moro 5, 00185, Rome, Italy*

[d] *Insect Behaviour Group, Dept. of Agriculture, Food and Environment, University of Pisa, Via del Borghetto 80, 56124, Pisa, Italy*

Received 10 December 2015; revised 29 March 2016; accepted 30 March 2016

Available online 12 July 2016

Abstract

Pest management tools should rely to proper identification of arthropod species, which are usually classified relying on morphological keys. However, the shortcomings and limitations of the conventional taxonomical identification methods highlighted need for new and simple methods of pest identification. In this research, DNA barcoding was used to identify 15 insect pests attacking horticultural crops in South India. Accurate phylogenetic information and evolutionary divergence data were supported and evidenced by various parameters, including the rates of substitution, nucleotide composition, genetic divergence, test of selection and saturation analysis. The null hypothesis of neutral selection was rejected in favour of the alternate hypothesis of purifying selection. NJ and ML trees were calculated and the individuals of the same species clustered together based on the CO1 sequence similarity, regardless of their collection site and geographic location. Overall, this study adds basic knowledge to molecular ecology of important insect pests attacking horticultural crops in South India.

Keywords: Insect pest; DNA barcoding; Molecular ecology; Nucleotide divergence; Phylogenetics

1. Introduction

A defining feature of the insect kingdom is their huge diversity in ecological and basic biological traits. Trophic interactions among plants, pests and pathogens pose economic and ecological threats to agricultural and wild land ecosystems [1]. Insect pests affecting succulent crops belonging to Solanaceae and Malvaceae are usually polyphagous. Their population outbreaks have an enormous potential to damage agricultural economy. Recognising the early signs of pests and diseases in order to deal with the problem is of crucial importance. Therefore, the accurate taxonomic identification is the pivotal issue in biological research, in order to allow the implementation of adequate measures to contend with species of medical

* Corresponding author.
E-mail address: karthikaconbio@gmail.com (P. Karthika).
Peer review under responsibility of University of Kerbala.

http://dx.doi.org/10.1016/j.kijoms.2016.03.006

or agricultural importance [2]. On the other hand, misidentifications could lead to ineffective control measures that potentially increase the impact caused by a particular pest species. Studies in biodiversity, systematics, community ecology and bio-monitoring also depend on proper taxonomic identification [3]. The standard method of utilization of morphological characters becomes challenging due to various factors such as phenotypical variations [4]. Detection of pests belonging to different groups is required for the protection of horticultural crops because they are prone to damage by a varied class of insects, including Hemiptera, Coleoptera and Lepidoptera [5].

In the last three decades, mitochondrial DNA has been extensively analysed [6] and proven to be an important tool in species delimitation as it possesses biological properties making it suitable as a marker for molecular biodiversity [7,8]. A major feature of DNA barcoding is that it allows prompt identification of pest young instars, as well as of fragmentary cuticular bodyparts. Partial DNA sequences of the mitochondrial gene such as Cytochrome c oxidase I (COI) and other molecular markers have been used to identify and discover new species (DNA barcoding). Several studies have shown that a 648 bp fragment of COI can be used as a DNA barcode to identify and distinguish between animal species [9,10]. Fragment size of COI has been shown to provide high resolution to identify cryptic species, thereby increasing taxonomy-based biodiversity estimates [11] and its usefulness has been confirmed for Coleoptera [12], Diptera [13,4], Ephemeroptera [14], Hemiptera [15], Hymenoptera [16] and Lepidoptera [17]. Nonetheless, DNA barcoding has proved to be a versatile tool with a variety of applications, for example, by facilitating the association between different developmental stages in insects [18] and the approach has also proved to be an effective auxiliary tool for forensic sciences [19], studies on feeding ecology [20] and habitat conservation initiatives [21]. Most importantly, DNA barcoding has proved to be especially useful in the study of taxonomically challenging taxa, where morphology-based identifications are frustrated due to cryptic diversity [22,23], or phenotypic plasticity [24]. Molecular phylogenies based on a large number of DNA markers have also revealed that species identification based on classical taxonomy alone may be misleading, with some species erroneously defined on the basis of associations with host plants [25,26].

Currently, more than 59,000 insect species have been described in India, of which only 4.6% barcodes have been generated from the known species, while the global record is about 16% of the described species [27]. In this research, DNA barcoding was used for identification of 15 insect pests attacking horticultural crops in South India. Our main goal was to use the mitochondrial sequence information from COI genes to shed light on molecular ecology of insect pests affecting the economically important vegetable crops. This will allow to evaluate the determinants or the various sequence parameters of accuracy, in order to study their divergences and phylogenetic relationships.

2. Materials and methods

2.1. Taxonomical collection and preservation

In South India, 26 insect species were mainly responsible of significant damage to important vegetable crops, like *Abelmoschus esculentus* (L.) Moench, *Lagenaria siceraria* (Molina), *Solanum melongena* (L), *Lablab purpureus* (L. Sweet) and *Trichosanthes cucumerina* (Linn.). Here, all pest samples were collected from three localities close to Coimbatore: Karamadai (11.16° N, 76.58° E), Pollachi (10. 40° N, 77. 01° E) and Annur (11.20 N° 77.23 E°) by hand-picking method and singly stored in clean glass vials. The study was conducted after obtaining permissions from the landowners. Insects belonging to adult and larval stages were identified by ZSI (Zoological Survey of India, Regional Research Stations, Chennai, India) relying to classic taxonomical keys. Samples were preserved in 75–95% ethanol (based on the nature of insects the percentage of ethanol should maintained for avoiding the decolouration) and kept at −20 °C in laboratory for further molecular investigations.

2.2. Molecular analysis, PCR and sequencing

Genomic DNA was isolated from thorax or whole body, depending on the insect size. The DNA was extracted by DNeasy kit (Qiagen, Germany) according to manufacturer's protocol. 1% Agarose Gel Electrophoresis (GENEI, Bangalore) was performed to detect the genomic DNA using Gel documentation (Medic Care, India). DNA amplification of COI gene was carried out by using ABI Thermocycler with the following primers for COI gene forward (LCO1490: 5′-GGTCAACAAATCATAAAGATATTG-3′) and reverse (HCO2198: 5′-TAAACTTCAGGGTGACCAAAAAATCA-3′) [28]. Amplification was performed in a total volume of 50 μl containing 4 μl of DNA template, 20 pM of each primers, 400 μM of dNTP and 0.4 μl of Taq DNA polymerase (Qiagen).

Thermo cycler conditions were as follows: 5 min at 95 °C for the initial phase, 35 cycles of 60 s at 95 °C for denaturation, 60 s at 52 °C for annealing, and 90 s at 72 °C for extension followed by 5 min at 72 °C for a final extension. The final PCR products were stored at −20 °C for further testing. The amplified products were resolved with 2% AGE and sequencing was done by using ABI 3500 XL Genetic Analyzer with manufacturer's protocol of Chromos Biotech, Pvt. Ltd., Bangalore, India. The sequences were trimmed and edited using Bio Edit v7.2.5 [29] and submitted to NCBI GenBank. Multiple sequence alignment was done with T-COFFEE.

2.3. General properties of sequences and genetic divergence

Similarity search for each sequence was performed using BLASTn. Fifty-four sequences (haplotypes and congeneric species) were retrieved from NCBI and used for comparisons (S1 Table 1). The genetic divergence was calculated using the Kimura-2-parameter (MEGA.v.6). Overall AT bias and nucleotide sequence was computed using DnaSp.v.5.1 [30] (Supplementary Online Material Table S1).

2.4. Rates of substitution and test of selection

The rate of transitions (TS) and transversions (TV) at the first, second and third codon positions were calculated and plotted against the F84 genetic using DAMBE 5.3.10 [31]. Test of sequence saturation was performed to estimate the transition/transversion versus the genetic distance (F84). Further investigation test for substantial saturation was checked by DAMBE [32,33]. Rate of synonymous (K_s) and non-synonymous substitutions (K_a) was calculated by Li93 method of DAMBE 5.3.10. The codon based Z-test of selection was done to test the type of selection. The probability of rejecting the null hypothesis of neutral evolution in favour of the alternate hypothesis (positive/purifying selection) was done based on the p-values ($p < 0.05$ are significant). Analyses were conducted using the Nei-Gojobori method [34] for sequence pairs and all evolutionary analyses were carried out using MEGA v.6 [35].

2.5. Phylogenetic analysis

The optimum substitution models were determined by the best-fit model test for the selection of model to be applied for sequences. The GTR + G model was selected from 24 different nucleotide substitution models for the 15 original sequences of the selected species and also for the haplotypes sequences (15 and 54 GenBank mined sequences based on the AIC (Akaike Information Criterion) and lowest BIC (Bayesian Information Criterion) values). The robustness of the clades of the tree was estimated using bootstrap analysis of 1000 replications with the elimination of all the codons containing gaps and missing data. The MP tree was constructed with the out-group of *Macrobrachium rosenbergii.* The tree was obtained by TBR (Tree Branch Swapping), the CI and RI and composite index were presented. The strength of the clades was assessed with the 10 random addition replicates and bootstrap analysis. Branch lengths were calculated by average pathway method and presented as the units of number of changes over the whole sequences. All the phylogenetic analysis was carried out in MEGA.v.6. The evolutionary history was inferred using the Neighbour-Joining method [36]. The optimal tree with the sum of branch length = 9.15978096 is shown. The evolutionary distances were computed using the Jukes-Cantor method [37] and are in the units of the number of base substitutions per site. The rate variation among sites was modelled with a gamma distribution (shape parameter = 1). The analysis involved 69 nucleotide sequences. There were a total of 1118 positions in the final dataset. Evolutionary analyses were conducted in MEGA6.

3. Results

A total of 520 individual insects belonging to 20 families of 4 orders (Coleoptera, Hemiptera, Lepidoptera and Diptera) were collected from *A. esculentus, L. siceraria, S. melongena, L. purpureus* and *T. cucumerina* (Table 1). From it, 15 pest species were chosen based on importance for Indian crops and the dataset included different developmental instars. Genomic DNA extracted from the tissue samples yielded >15 Kb and quality of the genomic DNA indicated appreciable amplification process with an approximate 700 bp size. The amplification success was higher in insects belonging to Coleoptera and Lepidoptera over Hemiptera. Similarity of the sequences was checked by using BLAST. The identity ranged between 82 and 99%. Among the 15 sequences studied here, 7 species *Aulacophora foveicollis* (Lucas), *Aulacophora cincta* (Fabricius), *Leptocentrus taurus* (Fabricius), *Cletus punctiger* (Dallas), *Heterorrhina elegans* (Fabricius), *Gametis versicolor* (Fabricius) and *Sphenarches caffer* (Zeller) are novel and

Table 1
Main insect pests of horticultural crops in South India, with the relative GenBank records referring to molecular taxonomical analyses.

Order	Family	Insect species*, location and GenBank accession numbers	Damaging instar	Host plant
Coleoptera ($n^a = 240$)	Coccinellidae	*Illeis cincta* (Fabricius)	Larva and adult	*Solanum melongena*
	Rutelidae	*Popillia cupricollis* (Hope)	Adult	*Solanum melongena*
	Curculionidae	*Myllocerus subfasciatus* (Guerin-Meneville); KJ183100, (A)	Adult	*Solanum melongena*
	Cetoniidae	*Gametis versicolor* (Fabricius); KJ559401, (A)	Adult	*Solanum melongena*
	Scarabaeidae	*Heterorrhina elegans* (Fabricius); KJ559404, (A)	Adult	*Solanum melongena*
	Curculionidae	*Myllocerus maculosus* (Desb.)	Adult	*Solanum melongena*
	Meloidae	*Mylabris phalerata* (Pallas)	Adult	*Abelmoschus esculentus*
	Coccinellidae	*Henosepilachna vigintioctopunctata* (Fabricius); KJ559395, (A)	Larva and adult	*Abelmoschus esculentus*
	Chrysomelidae	*Aulacophora foveicollis* (Lucas); KJ183105, (A)	Adult	*Lagenaria siceraria*
	Chrysomelidae	*Aulacophora cincta* (Fabricius); KJ183106, (A)	Adult	*Lagenaria siceraria*
	Chrysomelidae	*Aulacophora intermedia* (Jacoby)	Adult	*Lagenaria siceraria*
	Coccinellidae	*Menochilus sexmaculata* (Fabricius)	Adult	*Solanum melongena*
Hemiptera ($n^a = 180$)	Aleyrodidae	*Bemisia tabaci* (Gennadius)	Adult	*Solanum melongena*
	Pyrrhocoridae	*Dysdercus koenigii* (Fabricius); KJ183102, (A)	Adult	*Abelmoschus esculentus*
	Membracidae	*Leptocentrus taurus* (Fabricius); KJ183101, (A)	Adult	*Solanum melongena*
	Eurybrachidae	*Eurybrachys tomentosa* (Fabricius); KJ559400, (A)	Adult	*Lablab purpureus*
	Plataspididae	*Megacopta cribraria* (Fabricius); KJ559396, (A)	Adult	*Lablab purpureus*
	Coreidae	*Cletus punctiger* (Dallas); KJ183103, (A)	Adult	*Lablab purpureus*
	Aphididae	*Aphis craccivora* (Koch)	Adult	*Lablab purpureus*
	Pentatomidae	*Nezara viridula* (Linnaeus); KJ559399, (A)	Adult	*Lablab purpureus*
	Alydidae	*Riptortus linearis* (Fabricius); KJ559409	Adult	*Lablab purpureus*
Lepidoptera ($n^a = 80$)	Pterophoridae	*Sphenarches caffer*(Zeller); KJ183104	Adult	*Lagenariasiceraria*
	Crambidae	*Diaphania indica* (Saunders)	Larva and adult	*Lagenariasiceraria*
	Pyraloidea	*Maruca testulalis* (Geyer); KJ559398, (A)	Larva and adult	*Lablab purpureus*
	Crambidae	*Leucinodes orbonalis* (Guenée)	Larva and adult	*Solanum melongena*
Diptera ($n^a = 20$)	Tephritidae	*Bactrocera dorsalis* (Hendel)	Adult	*Trichosanthes cucumerina*

*The n. of individuals sampled for each species is 20 (Karamadai (K)- 11.16° N, 76.58° E = 6, Pollachi (P)-10. 40° N, 77. 01° E = 7, Annur (A)-11.20 N° 77.23 E° = 7), in all the three study sites K, P and A.
[a] Number of individuals

represented first time records submitted to the NCBI GenBank database. Multiple sequence alignment showed 80% identical entries and highly similar nucleotide regions. A total of 15 species were sequenced and the aligned sequences possessed a range between 528 and 721bp of the COI region and it was authenticated by GenBank with the following accession numbers KJ183100- KJ183106, KJ559395, KJ559396, KJ559398, KJ559399-KJ559401, KJ559404 and KJ559409.

The analysed set of sequences had 372 variable sites of which 221 were parsimony informative (including the out-group *M. rosenbergii*) with the consistency index (CI) of 0.442 and retention indices (RI) of 0.32. Generally, like other protein coding genes majority of the substitutions occurred in the 3rd codon position. Estimation of transition and transversion substitutions was studied and it was high in A and T nucleotides with a maximum of 41.03% in 3rd position; 29.13% and 30.09% in 2nd and 1st positions respectively. In addition to the above, probable substitutions occurred in the synonymous sites of the nucleotide sequences. The base composition of the COI fragment varied among the species but it was commonly demonstrated with an overall AT bias of 67.27 and GC of 32.73. The A + T contents of the 1st, 2nd and 3rd codons were 58.87, 55.24 and 82.08% respectively. The three codon positions of the CO1 gene sequences of the present study were tested for saturation by plotting the number of transitional (TS) and transversional substitutions (TV) against the F84 for all the three codons and were represented as 3-D plot (Fig. 1). The plot showed that the TS and TV for the first and third codon position increased gradually along the F84 distance. However the substitutions in the second codon position was less. Exact saturation of the TV and TS did not occur in all the three codon positions.

The test of substitution saturation of the COI sequences transitions and transversions showed that the number of transversions (TV) outnumbered transitions (TS) (Supplementary Online Material Figure S1). If sequences exhibit substitution saturation, they possess less phylogenetic information and therefore could not be able to reveal deeper insights for phylogenetic

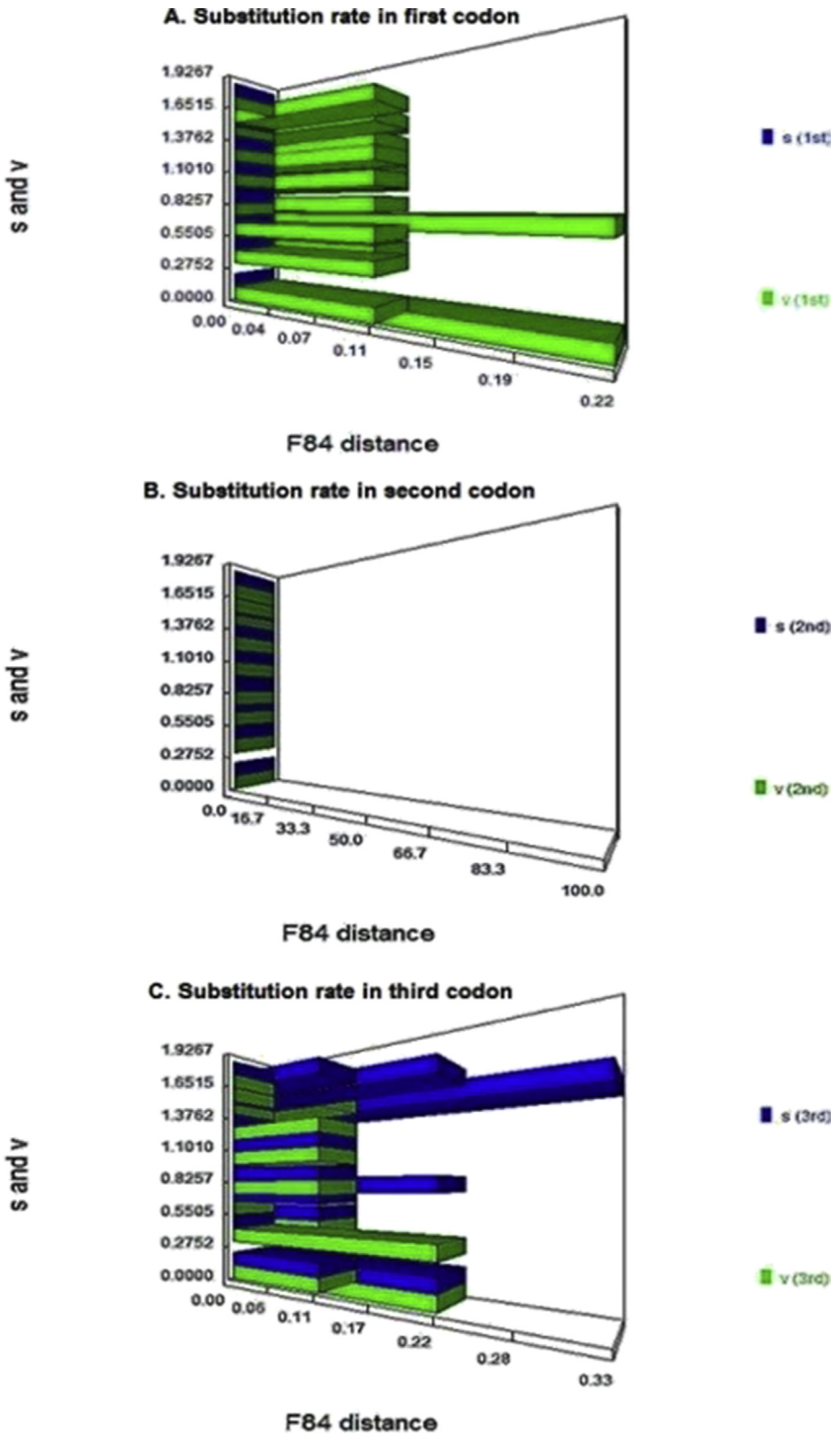

Fig. 1. Number of substitutions (s = transitions, v = transversions) in y-axis against F84 genetic distance in x-axis for each codon position, represented as A, B and C.

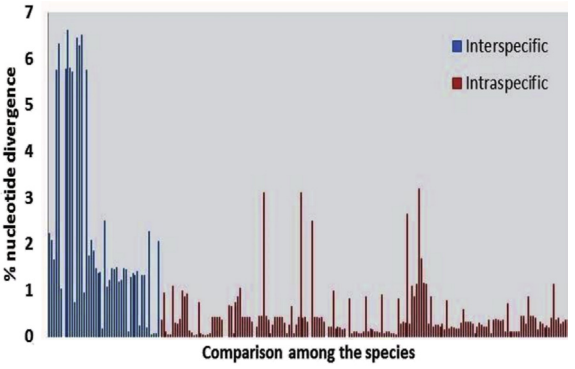

Fig. 2. Distribution of genetic distance within and between the pest species; the entries for intraspecific divergence are approximately <1.5%, higher divergence is clearly indicated in the interspecific over intraspecifc comparisons among species; the threshold level is estimated as 0.4% among the species.

signal. However, the test of substantial saturation revealed that the value of Iss is lower than the Iss.c [0.6711 (Iss) < 0.743 (Iss.c)]. Hence, it was evident that the selected sequences of the current study did not undergo substantial saturation.

The genetic divergence among the selected species belonging to different orders reported in Fig. 2 and all the species exhibit discriminative values of inter and intraspecific divergence. The congeneric and conspecific divergences were calculated based on the availability of the sequences in NCBI database. Accordingly, the average congeneric divergence of *Myllocerus, Dysdercus, Cletus, Aulacophora* and *Maruca* were 0.397, 0.437, 0.388, 0.323 and 0.42 respectively. The average congeneric divergence was recorded as 0.393. The comparison of the 15 sequences of this study with the existing congeneric sequences exhibited high divergence among the *Myllocerus subfasciatus* (India) and *Myllocerus variabilis* (Korea) with 0.30159; *Dysdercus koenigii* (India) and *Dysdercus evanescens* (India) with 0.63226; *A. foveicollis* (India) and *Aulacophora indica* (Taiwan) 0.349; *C. punctiger* (India) and *Cletus trigonus* (China) 0.32809; *Maruca testulalis* (India) and *Maruca vitrata* (South Africa) 0.391. It is noteworthy that maximum congeneric divergence was exhibited between the congenerics of *Dysdercus* from the same geographical location (India). Since no other haplotypes for 13 pest species were available in the NCBI, the conspecific divergence was calculated for available haplotypes of *Henosepilachna vigintiopunctata* and *Nezara viridula.* The overall divergence of the conspecifics was 0.0420. In the case of *H. vigintiopunctata*, high conspecific divergence was found with the haplotypes from Indonesia (0.093). In *N. viridula* 27 haplotypes were retrieved from 7 different geographical locations, including India. Obviously, the lowest conspecific divergence was noted in the Indian haplotypes (0.0136) and the highest in the haplotypes of Slovenia (0.0622). The effectiveness of the pairwise comparison and divergence showed that the congeneric divergence was nine times higher than the conspecifics. In conspecifics, most of the values were <2%. The threshold level of nucleotide divergence was 0.420%. In relevance to the 10× rule (Hebert et al., 2003), most of the cases were well resolved up to species level in the intraspecies comparisons while majority of the haplotype sequences of the major pests were unavailable and there is a wide lacuna in exploring these particular agriculture pests.

Synonymous and non-synonymous substitutions were represented as K_a and K_s in S1 Fig. 2. K_s were found higher than K_a (Supplementary Online Material Figure S2). The Z-test of selection was performed to trace the evolutionary lineage among the haplotypes and showed in Fig. 3. The estimation of positive selection revealed that the test statistic (dN- dS) possessed maximum negative values and the p-values were >0.01. On contrary, the estimation of purifying selection had maximum positive values for the test statistic (dS- dN) and the maximum p-values were <0.01. Hence, the probability (P) of rejecting null hypothesis in strict neutrality in favour of the alternate hypothesis of purifying selection was true for the evolutionary selection of the sequences of the selected pests.

In the ML analysis, the best-fit model of the nucleotide substitution was evaluated based on the AICc value (Akaike Information Criterion, corrected) 9350.26 and BIC scores (Bayesian Information Criterion) 9610.05. The GTR + G model was selected for ML analysis of the 15 sequences representing the partial CO1 gene sequence. Bootstrapping of the ML analysis (1000 replications) was implemented in NNI ML heuristic method with highest log likelihood (−4413.51) shown in Supplementary Online Material Figure S3. The CO1 sequences of the dataset were represented in a phylogenetic tree with distinctive clusters based on the nucleotide divergence among the species. The upper clade was formed with a group of species belonging to Hemiptera with a branch length of 0.027 whereas the subsequent clade below was constituted by the major Coleoptera species. The branch length between these major clusters was

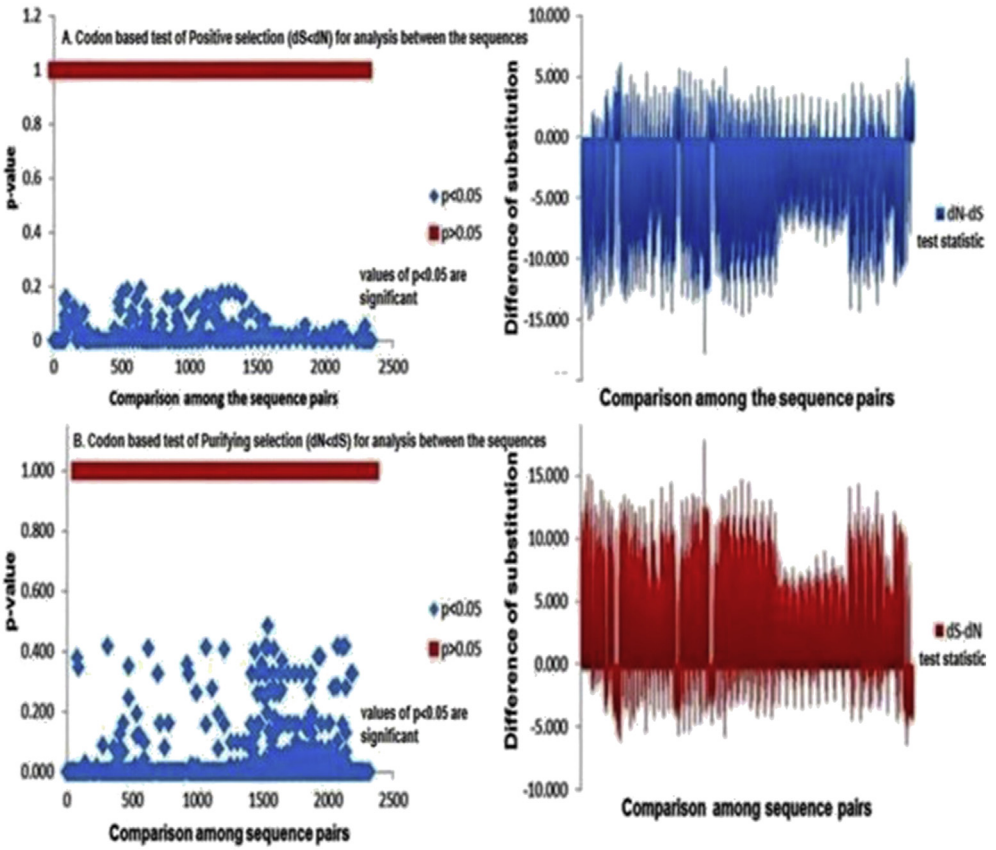

Fig. 3. Z-test of selection inferred by sequence pairs, Scatter plots show the p-values marked to each type of selection separately in which $p < 0.05$ are significant(left), the graphs show the test statistic (dN-dS) and (dS-dN) for positive and purifying selection respectively (right).

minimum. Whereas species representing Lepidoptera occurred separately with distinctive branch length of 0.173 (*S. caffer*) and *M. testulalis* had minimum branch length from *S. caffer. M. subfasciatus* and *H. elegans* formed significant discrete branches with maximum branch length 11.776 and 0.4479, respectively.

The phylogram for the haplotypes and the sequences of the present study was constructed using the GTR + G, which was selected by the best fit model test based on the AIC (5768.09) and BIC (6812.33) scores. Comparison of the COI sequences along with their haplotypes yielded distinctive polyphyletic tree with typical conspecific and congeneric clusters shown in Fig. 4. *Henosepilachna vigintioctopunctata, D. koenigii* and *N. viridula* formed major conspecific clusters and that of *Aulacophora, Myllocerus* and *Maruca* formed significant congeneric groups. Meanwhile, *H. elegans*, *L. taurus*, *G. versicolor*, *Eurybrachys tomentosa* and *Megacopta cribraria* formed separate clades because of high variability among the haplotypes and the original sequences of this study. The *C. punctiger and Cletus schmiditi* clustered together, whereas the other congenerics of *Cletus* occurred in distant separate clades. Since the nucleotide sequence of the seven species were new in GenBank, their haplotypes were not available, and this perhaps be the possible reason for their occurrence with other species due to minimum nucleotide similarity. The MP tree was obtained using the Tree-Bisection-Regrafting (TBR) algorithm with search level 1 in which the initial trees were obtained by the random addition of sequences (10 replicates) with most parsimonious tree with length = 1150. The MP tree was constructed with *M. rosenbergii* as the out-group had significant bootstrap values near the braches shown in Supplementary Online Material Figure S4. *H. elegans* had a higher divergence (branch length = 207.93) which was greater than out group. Similarly, the branch length of other species *M. subfasciatus* (80.952) >*L. taurus* (70.140) >*M. rosenbergii* (64.13) (see also Supplementary Online Material Figure S4).

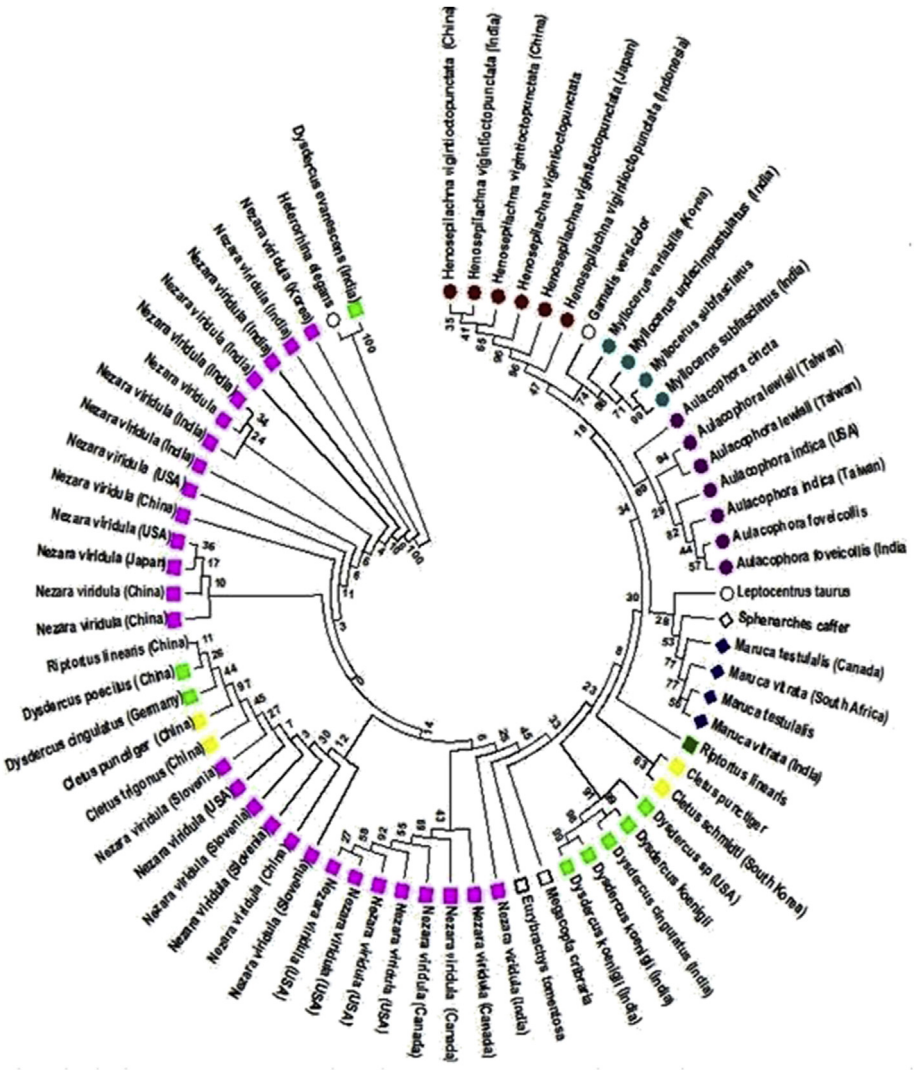

Fig. 4. NJ tree formed by GTR + G model with the hapolotypes along with the selected species of this study in the paraphyletic clade with initial tree(s) for the heuristic search obtained automatically by applying Neighbour-Join and Bio NJ algorithms.

4. Discussion

Molecular ecology research on insect pests of agricultural importance may lead to identifying new species [38], biotypes [39], cryptic species [11] and haplotypes [40], which are difficult to study relying only to morphological characters. Sustenance of crop productivity and nutritional security encounters high challenges through direct and indirect damages caused insect pests [41] In this scenario, while classical taxonomy has its own strengths, molecular identification employing COI barcoding has the added advantage of not being limited by polymorphism, sexual form (asexual/sexual) and life stages of the target species [42]. All the species employed in the present study were differentiated on the basis of COI gene, however amplification using the universal primers. The yield of DNA had better results in the species belonging to Coleoptera and Lepidoptera, than that of Hemiptera. Similarly the universal primers used in aphid (Heteroptera) DNA barcodes also amplified barcode sequences from parasitoid species within the host aphids, which indicated serious problems in the use of universal primers for aphid barcoding [43].

DNA sequences are characterized by summary statistics like length and base composition. Prior to comparison of nucleotide sequences in phylogenetic analysis, several additional parameters like overall rate of nucleotide substitution, ratio of two specific instantaneous rates of substitution rate at which

transitions and transversions occur and the rate variation among sites play a significant role and are necessary for accurate reconstruction of phylogeny [44,4]. The present study was based on the molecular identification of pests in economically important vegetable crops and it comprised of 15 nucleotide sequences of insect pests which were identified by COI. Results indicated that the COI-based pest identification was extremely effective for insects because all the species were accurately and successfully identified based on the COI marker profile. Most of the phylogenetic information has been derived from mitochondrial DNA variations [45] and recently DNA sequence data have been employed successfully to elucidate the relationships of many groups of insect species at generic level [46,47]. The composition of the mitochondrial sequence of the COI gene in the present study was expectedly AT biased and this was generally observed in several previous studies [48]. In general, the frequency of transitional substitutions is known to be higher than transversional substitutions in the genome [49], saturation possibly occur when the plot shows no further increase in the observed number of transitions despite increasing genetic distances. Thus, saturation of transitions at high levels of sequence divergence indicates saturation in the data. However, in our experiments the rate of transversions was higher than that of transitions, these substitutions perhaps lead to substantial saturation of the sequences. Analysing the transitions and transversions separately in this manner showed that transversions contain stronger phylogenetic signal than transitions, and are capable of masking the distorted signals coming from the saturated transitions sites that may be misleading. In addition the plot of TV/TS in the 3rd codon position against the F84 distance also conveyed that transversions were maximum. Highly divergent datasets are still usable, as the phylogenetic information is of not completely lost and may be retained using sites that have experienced more transversions [44] However, the sequences of the study correspond to individuals from highly diverged groups such as Coleoptera, Hemiptera and Lepidoptera. Therefore it could be a reason for the wide differences in substitutions. Indeed, it was evident that no substantial saturation occurred since Iss < Iss.c This adequate result confirmed that the sequences did not undergo substantial saturation and the sequences could be used for inferring phylogenetic information and reconstruction. Similar results were seen in *Toxoptera* genus, where the transitions and transversions had a linear relationship and demonstrated that they were not saturated [50].

DNA barcoding studies on insects and invertebrates have shown maximum intraspecific variation ranging from 3 to 3.9% [51]. According to earlier studies [11], sequence thresholds of 10X, the mean intraspecific mt COI variation for group under study as an approximate method of screening for potentially different (cryptic) species or biotypes. The genetic divergence was computed using K2P.The threshold value of the CO1 sequences was found to be 0.413%, notably with the interspecific divergence of 3.93% and intraspecific divergence of 0.41%. In 96.4% of the cases, analysis of evolutionary divergence using the K2P model and the 3% cut-off criteria suggested for species level divergence [8], identified clades that were in concordance with recognized taxonomic units based on morphological characters. According to 10X rule the percentage of nucleotide divergence between the intraspecies should be less than 3% and that of interspecies should be more than 3%. Hence the sequences analysed in the present study exhibited high inter species variability on the basis of nucleotide sequences. Therefore the intraspecific divergence was higher enough to discriminate between the individuals. The divergences among inter and intra species was closely associated. As expected, the conspecific values (0.0420) were lower than the congeneric (0.3930). The values of the sequence divergence varied widely across the taxa, similarly putative species like *N. viridula* exhibited high levels of intraspecific divergence. It is noteworthy to mention that in terms of conspecific comparisons, the haplotypes belonged to wide geographical localities therefore high variations were exhibited. Cytochrome oxidase I (COI) sequences were shown to be appropriate for intraspecific analysis of the observed high degree of polymorphism. Furthermore, CO1 sequences have been used in some studies to address similar problems on a comparable geographic range, and using the same marker might facilitate comparisons in *Bactrocera dorsalis* (Diptera: Tephritidae) [52]. In a study of Cuban fresh water fishes, the level of divergence among congeneric species was about 10 times higher than among conspecifics, and divergence levels between confamiliars were about two times higher than congenerics with the mean conspecific, congeneric and confamiliar genetic distances were 0.4%, 8% and 20.4% respectively [53]. While in the case of other arthropods like *Toxoptera*, the interspecific divergences were 7.37%–9.19% which were greater than intraspecific divergences, with a mean of 8.29% [50] This value nears the mean level of congeneric divergence value (11.3%) in large phyla [8], and is similar to interspecies genetic distances

between *Schizaphis graminum* and *S. rotundiventris*, which range from 8.44% to 9.99% [54]. Similar studies indicated that divergence of 13, 320 species pairs ranged from a low of 0.0% to a high of 53.7%, while most pairs (79%) showed greater than 8% sequence divergence. In researches on the evolutionary divergence of DNA sequences, it is often required to estimate the numbers of synonymous (silent) and non-synonymous (amino acid altering) nucleotide substitutions separately. Since the rate of synonymous substitution is much higher than that of non-synonymous substitution and is similar for many different genes, synonymous substitutions may be used as a molecular clock for dating the evolutionary time of closely related species [55,56]. In parallel to the above, the nucleotide sequences of the present study were estimated for the synonymous (K_s) and non-synonymous (K_a) and obviously the rate of K_s was higher in overall comparisons. On contrary, the earlier studies on the comparativeness of substitutions in the mitochondrial cyt b and ND2 genes of island birds was reported in which, among all cases more non-synonymous substitutions occurred on the island branch, but synonymous substitutions showed no significant bias thus they experience to undergo positive selection [57]. Similarly, the evidence of genetic drift among bacterial endosymbionts of aphids also possessed high ratios of nonsynonymous to synonymous substitutions in *Buchnera* loci relative to the enteric bacteria, which strongly suggested the accumulation of nonsynonymous substitutions in endosymbiont lineages. These combined, results suggested that the endosymbiont populations exhibited decreased efficacy of purifying selection [58].

Typically K_a/K_s defined the ratio of number of non-synonymous substitutions per non-synonymous site (K_a) to the number of synonymous substitution per synonymous site (K_s), these serve as an important factor in determining the evolution selection pressure of an organism. The mitochondrial COI gene appears to be among the most conservative protein-coding genes in the mitochondrial genome of animals [59], if a region is said to be highly conserved between the species it means that the important proteins are coded by that region and it provides very vital function for the organism to survive, therefore if that particular region is damaged or lethally mutated the organism may not survive and cannot pass on the necessary changes to offspring. Such genes are preferable for the understanding of natural selection of the sequences Therefore as an effective and supportive parameter, K_a/K_s was calculated for the mitochondrial nucleotide sequences of this research. Selection is a crucial parameter to describe the evolutionary process in phylogenomics studies. Estimation of non-synonymous (K_a) and synonymous (K_s) substitution rates is of great significance in reconstructing phylogeny and understanding evolutionary dynamics of protein-coding sequences across closely related and yet diverged species [60]. Previous studies [61] suggested that usage of statistical methods in evolutionary genetics could evaluate the strength of selection operating on individual codons over particular branches or regions of a phylogenetic tree. The null hypothesis of neutral selection was rejected since the p-values <0.05 which were considered to be significant. Further, the test statistic dN-dS had a majority of negative values, which clearly indicated the absence of positive selection. Hence it could be evidently noted that synonymous substitutions accumulated highly in the analysed sequences of the current study and since $dS > dN$, it interprets for the purifying selection.

Molecular identification was done for several pests worldwide, in *Orius* (Hemiptera: Anthocoridae) [62] and potato flea beetles (Coleoptera: Chrysomelidae) [63]. Moreover, discrimination of aphids of 32 species collected in various host plants in South India was also reported [41]. The present study was completely outlined on a group of important insect pests attacking field crops in South India; hence it was a varied and wide combination of individuals from distantly related taxa. This could be an essential breakthrough and unique difference from earlier studies. This investigation of COI barcoding could potentially be applied in agricultural and horticultural researches to rapid identification of pests. The phylogenetic signal is a direct function of the length of the branch (in units of the expected number of substitutions per site), which sheds light the evolutionary relationship [64]. MP analysis of the mitochondrial DNA has been used in the phylogenetic relationships among the pests but the relationships among the taxa could not be well resolved because in majority of the MP reconstructions, the chosen outgroup species would retain higher divergence.

Previous studies have analysed the divergence among the closest set of individuals either belonging to a single genera, family or species level. On contrary, the selection and molecular identification of individuals for the present study was based on level of damages caused by them to selected vegetable crops, therefore it cannot be narrowed down to a single criterion like family or genera. This could be a strong reason for the unique results of the study and it is definitely distinctive from the studies done so far.

Moreover, the selection of organisms for molecular identification was truly based on the field studies which provided basic information on the highly damage causing pests. Secondly, several taxonomical ambiguities were faced during the collection and identification stages, chiefly in the larval pests and also to a certain extent in the species level. Moreover, all the species were accurately identified with maximum identity values, the properties of the DNA sequences were precisely studied and subsequently utilized for inferring the evolutionary data and the genetic divergence was appropriately congruent. Therefore the present study had potentially provided deeper insights of knowledge of pests in the major vegetable crops of India based on their nucleotide sequences. Future research based on employing of species-specific primers is needed to yield qualitative amplified products upon which the success of sequencing is dependent. The inclusion of one or more genetic markers and the comparison of the data among them would provide a better, effective, deeper insight for understanding the evolutionary relationships among the organisms.

5. Conclusions

In this study, the analysis of COI sequences exhibited high divergence among the selected insect pests. We highlighted essential information on important insect pests attacking vegetable crops in South India and it also provided information on their association in the host plants due to their polyphagous feeding behaviour. DNA barcoding using COI genes could be an effective method for screening insect pests and to shed light on their genetic variations in addition the integration of traditional taxonomy [65]. Overall, our findings contribute to a better understanding of the identification of pests by COI genes and aid in improving Integrated Pest Management in Asian countries.

Acknowledgements

The authors gratefully acknowledge the Department of Science and Technology-India for providing financial support through DST-INSPIRE IF130010 (GOVERNMENT OF INDIA).

Appendix A. Supplementary data

Supplementary data related to this article can be found at http://dx.doi.org/10.1016/j.kijoms.2016.03.006.

References

[1] D. Pimentel, L. Lach, R. Zuniga, D. Morrison, Environmental and economic costs of nonindigenous species in the United States, Bioscience 50 (2000) 53–65. http://www.bioone.org/doi/abs/10.1641/0006-3568(2000)050[0053:EAECON]2.3.CO;2.

[2] D.R. Miller, A.Y. Rossman, Systematics, biodiversity and agriculture, Bioscience 45 (1995) 680, http://dx.doi.org/10.2307/1312673.

[3] J. Rivera, D.C. Currie, Identification of Nearctic black flies using DNA barcodes (Diptera: Simuliidae), Mol. Ecol. Resour. 9 (2009) 224–236, http://dx.doi.org/10.1111/j.1755-0998.2009.02648.x.

[4] K. Murugan, C. Vadivalagan, P. Karthika, C. Panneerselvam, M. Paulpandi, J. Subramaniam, H. Wei, M.S. Alsalhi, S. Devanesan, M. Nicoletti, R. Paramasivan, DNA barcoding and molecular evolution of mosquito vectors of medical and veterinary importance, Parasitol. Res. (2015), http://dx.doi.org/10.1007/s00436-015-4726-2.

[5] Denholm, M. Cahill, T.J. Dennehy, A.R. Horowitz, Challenges with managing insecticide resistance in agricultural pests, exemplisfied by the whitefly Bemisia tabaci, Philos. Trans. R. Soc. L. B Biol. Sci. 353 (1998) 1757–1767, http://dx.doi.org/10.1098/rstb.1998.0328.

[6] J.W.O. Ballard, M.C. Whitlock, The incomplete natural history of mitochondria, Mol. Ecol. 13 (2004) 729–744, http://dx.doi.org/10.1046/j.1365-294X.2003.02063.x.

[7] W.S. Moore, Inferring phylogenies from mtDNA variation-mitochondrial-gene trees versus nuclear-gene trees, Evolution 49 (1995) 718–726.

[8] P.D.N. Hebert, A. Cywinska, S.L. Ball, J.R. deWaard, Biological identifications through DNA barcodes, Proc. Biol. Sci. 270 (2003) 313–321, http://dx.doi.org/10.1098/rspb.2002.2218.

[9] P.D.N. Hebert, M.Y. Stoeckle, T.S. Zemlak, C.M. Francis, Identification of birds through DNA barcodes, PLoS Biol. 2 (2004a) e312, http://dx.doi.org/10.1371/journal.pbio.0020312.

[10] R.D. Ward, T.S. Zemlak, B.H. Innes, P.R. Last, P.D.N. Hebert, DNA barcoding Australia's fish species, Philos. Trans. R. Soc. Lond. B. Biol. Sci. 360 (2005) 1847–1857, http://dx.doi.org/10.1098/rstb.2005.1716.

[11] P.D.N. Hebert, E.H. Penton, J.M. Burns, D.H. Janzen, W. Hallwachs, Ten species in one: DNA barcoding reveals cryptic species in the neotropical skipper butterfly Astraptes fulgerator, Proc. Natl. Acad. Sci. U. S. A. 101 (2004) 14812–14817, http://dx.doi.org/10.1073/pnas.0406166101.

[12] R.A.B. Leschen, I. Lobl, Demography of coleopterists and their thoughts on DNA barcoding and the phylocode, with commentary, Coleopt. Bull. 59 (2005) 284–292, http://dx.doi.org/10.1649/850.1.

[13] S.J. Scheffer, M.L. Lewis, R.C. Joshi, DNA barcoding applied to invasive Leafminers (Diptera: Agromyzidae) in the Philippines, Ann. Entomol. Soc. Am. 99 (2006) 204–210 doi:10.1603/0013-8746(2006)099[0204:DBATIL]2.0.CO;2.

[14] S.L. Ball, P.D.N. Hebert, S.K. Burian, J.M. Webb, Biological identifications of mayflies (Ephemeroptera) using DNA barcodes, J. North Am. Benthol. Soc. 24 (2005) 508 doi:10.1899/0887-3593(2005)024＼[0508:BIOMEU＼]2.0.CO;2.

[15] W. Lee, H. Kim, J. Lim, H.-R. Choi, Y. Kim, Y.-S. Kim, et al., Barcoding aphids (Hemiptera: Aphididae) of the Korean Peninsula: updating the global data set, Mol. Ecol. Resour. 11 (2011) 32–37, http://dx.doi.org/10.1111/j.1755-0998.2010.02877.x.

[16] M.A. Smith, J.J. Rodriguez, J.B. Whitfield, A.R. Deans, D.H. Janzen, W. Hallwachs, et al., Extreme diversity of tropical parasitoid wasps exposed by iterative integration of natural history, DNA barcoding, morphology, and collections, Proc. Natl. Acad. Sci. U. S. A. 105 (2008) 12359−12364, http://dx.doi.org/10.1073/pnas.0805319105.

[17] M. Hajibabaei, D.H. Janzen, J.M. Burns, W. Hallwachs, P.D.N. Hebert, DNA barcodes distinguish species of tropical Lepidoptera, Proc. Natl. Acad. Sci. U. S. A. 103 (2006) 968−971, http://dx.doi.org/10.1073/pnas.0510466103.

[18] D. Ahrens, M.T. Monaghan, A.P. Vogler, DNA-based taxonomy for associating adults and larvae in multi-species assemblages of chafers (Coleoptera: Scarabaeidae), Mol. Phylogenet. Evol. 44 (2007) 436−449, http://dx.doi.org/10.1016/j.ympev.2007.02.024.

[19] N. Dawnay, R. Ogden, R. McEwing, G.R. Carvalho, R.S. Thorpe, Validation of the barcoding gene COI for use in forensic genetic species identification, Forensic Sci. Int. 173 (2007) 1−6, http://dx.doi.org/10.1016/j.forsciint.2006.09.013.

[20] A.K. Kuusk, N. Agustí, Group-specific primers for DNA-based detection of springtails (Hexapoda: Collembola) within predator gut contents, Mol. Ecol. Resour. 8 (2008) 678−681, http://dx.doi.org/10.1111/j.1471-8286.2007.02046.x.

[21] R.D. Ward, B.H. Holmes, W.T. White, P.R. Last, DNA barcoding Australasian chondrichthyans: results and potential uses in conservation, Mar. Freshw. Res. 59 (2008) 57−71, http://dx.doi.org/10.1071/MF07148.

[22] J.D.S. Witt, D.L. Threloff, P.D.N. Hebert, DNA barcoding reveals extraordinary cryptic diversity in an amphipod genus: implications for desert spring conservation, Mol. Ecol. 15 (2006) 3073−3082, http://dx.doi.org/10.1111/j.1365-294X.2006.02999.x.

[23] A. Yassin, P. Capy, L. Madi-Ravazzi, D. Ogereau, J.R. David, DNA barcode discovers two cryptic species and two geographical radiations in the invasive drosophilid Zaprionus indianus, Mol. Ecol. Resour. 8 (2008) 491−501, http://dx.doi.org/10.1111/j.1471-8286.2007.02020.x.

[24] S.J. Adamowicz, P.D.N. Hebert, M.C. Marinone, Species diversity and endemism in the Daphnia of Argentina: a genetic investigation, Zool. J. Linn. Soc. 140 (2004) 171−205, http://dx.doi.org/10.1111/j.1096-3642.2003.00089.x.

[25] A. Coeur d'acier, Armelle, G. Cocuzza, E. Jousselin, V. Cavalieri, S. Barbagallo, Molecular phylogeny and systematic in the genus Brachycaudus (Homoptera: Aphididae): insights from a combined analysis of nuclear and mitochondrial genes, Zool. Scr. 37 (2008) 175−193.

[26] E. Jousselin, Y. Desdevises, A. Coeur d'acier, Fine-scale cospeciation between *Brachycaudus* and *Buchnera aphidicola*: bacterial genome helps define species and evolutionary relationships in aphids, Proc. Biol. Sci. 276 (2009) 187−196, http://dx.doi.org/10.1098/rspb.2008.0679.

[27] S.K. Jalali, Rakshit Ojha, T. Venkatesan, DNA Barcoding for Identification of Agriculturally Important Insects. New Horizons in Insect Science: towards Sustainable Pest Management, Springer, India, 2015, pp. 13−23.

[28] O. Folmer, M. Black, W. Hoeh, R. Lutz, R. Vrijenhoek, DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates, Mol. Mar. Biol. Biotechnol. 3 (1994) 294−299, http://dx.doi.org/10.1371/journal.pone.0013102.

[29] T. Hall, Bio Edit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT, Nucleic Acids Symp. Ser. 41 (1999) 95−98 doi:citeulike-article-id:691774.

[30] P. Librado, J. Rozas, DnaSP v5: a software for comprehensive analysis of DNA polymorphism data, Bioinformatics 25 (2009) 1451−1452, http://dx.doi.org/10.1093/bioinformatics/btp187.

[31] X. Xia, DAMBE5: a comprehensive software package for data analysis in molecular biology and evolution, Mol. Biol. Evol. 30 (2013) 1720−1728, http://dx.doi.org/10.1093/molbev/mst064.

[32] X.H. Xia, Z. Xie, M. Salemi, L. Chen, Y. Wang, An index of substitution saturation and its application, Mol. Phylogenet. Evol. 26 (2003) 1−7, http://dx.doi.org/10.1016/s1055-7903(02)00326-3.

[33] X. Xia, P. Lemey, Assessing Substitution Saturation with DAMBE, in: The Phylogenetic Handbook: a Practical Approach to DNA and Protein Phylogeny, 2, 2009, pp. 615−630.

[34] M. Nei, T. Gojobori, Simple methods for estimating the numbers of synonymous and nonsynonymous nucleotide substitutions, Mol. Biol. Evol. 3 (1986) 418−426.

[35] K. Tamura, G. Stecher, D. Peterson, A. Filipski, S. Kumar, MEGA6: molecular evolutionary genetics analysis version 6.0, Mol. Biol. Evol. 30 (2013) 2725−2729, http://dx.doi.org/10.1093/molbev/mst197.

[36] N. Saitou, M. Nei, The neighbor-joining method: a new method for reconstructing phylogenetic trees, Mol. Biol. Evol. 4 (1987) 406−425 doi:citeulike-article-id:93683.

[37] T.H. Jukes, C.R. Cantor, Evolution of protein molecules, Mamm. Protein Metab. (1969) 21−123 doi:citeulike-article-id:768582.

[38] S.L. Ball, K.F. Armstrong, DNA barcodes for insect pest identification: a test case with tussock moths (Lepidoptera : Lymantriidae), Can. J. For. Res. Can. Rech. For 36 (2006) 337−350, http://dx.doi.org/10.1139/X05-276.

[39] T.M. Perring, The Bemisia tabaci species complex, Crop Prot. 20 (2001) 725−737, http://dx.doi.org/10.1016/S0261-2194(01)00109-0.

[40] S. Toda, T. Murai, Phylogenetic analysis based on mitochondrial COI gene sequences in *Thrips tabaci* Lindeman (Thysanoptera : Thripidae) in relation to reproductive forms and geographic distribution, Appl. Entomol. Zool. 42 (2007) 309−316, http://dx.doi.org/10.1303/aez.2007.309.

[41] K.B. Rebijith, R. Asokan, N.K.K. Kumar, V. Krishna, B.N. Chaitanya, V.V. Ramamurthy, DNA barcoding and elucidation of cryptic aphid species (Hemiptera: Aphididae) in India, Bull. Entomol. Res. 103 (2013) 601−610, http://dx.doi.org/10.1017/S0007485313000278.

[42] R. Asokan, K.B. Rebijith, S.K. Singh, A.S. Sidhu, S. Siddharthan, P.K. Karanth, et al., Molecular identification and phylogeny of Bactrocera species (Diptera: Tephritidae), Fla. Entomol. 94 (2011) 1026−1035, http://dx.doi.org/10.1653/024.094.0441.

[43] W. Lee, S. Lee, Unexpected problem in aphid DNA barcoding by universal primers, Entomol. Sci. 15 (2012) 121−126, http://dx.doi.org/10.1111/j.1479-8298.2011.00473.x.

[44] B. Dwivedi, S.R. Gadagkar, The impact of sequence parameter values on phylogenetic accuracy, Biol. Med. 1 (2009) 50−62.

[45] J.C. Avise, Phylogeography: the History and Formation of Species, Harvard university press, 2000.

[46] J.Y. Wang, J.F. Guo, D.Y. Zhao, C. Wang, Y. Zhang, H.Z. Wang, Y. Wu, The genetic diversity and phylogenetic relationship among pig breeds of Shandong Province based on mtDNA CytB gene, Sci. Agric. Sin. 42 (2009) 1761−1767.

[47] L. Ruo-Yu, Y. Gong-She, L. Chu-Zhao, The genetic diversity of mtDNA D-loop and the origin of Chinese goats, Yi Chuan Xue Bao 33 (2006) 420−428, http://dx.doi.org/10.1016/S0379-4172(06)60069-3.

[48] M. Zhang, T. Cao, R. Zhang, Y. Guo, Y. Duan, E. Ma, Phylogeny of Apaturinae butterflies (Lepidoptera: Nymphalidae) based on mitochondrial cytochrome oxidase I gene, J. Genet. Genomics 34 (2007) 812−823 doi:S1673-8527(07)60092-7 [pii]\r10.1016/S1673-8527(07)60092-7.

[49] J. Wakeley, The excess of transitions among nucleotide substitutions: new methods of estimating transition bias underscore its significance, Trends Ecol. Evol. 11 (1996) 158−163, http://dx.doi.org/10.1016/0169-5347(96)10009-4.

[50] J.F. Wang, G.X. Qiao, DNA barcoding of genus Toxoptera Koch (Hemiptera: Aphididae): identification and molecular phylogeny inferred from mitochondrial COI sequences, Insect Sci. 16 (2009) 475−484.

[51] M.E. Carew, V. Pettigrove, R.L. Cox, A.A. Hoffmann, DNA identification of urban Tanytarsini chironomids (Diptera : Chironomidae), J. North Am. Benthol. Soc. 26 (2007) 587−600, http://dx.doi.org/10.1899/06-120.1.

[52] S. Nakahara, M. Muraji, PCR-RFLP analysis of *Bactrocera dorsalis* (Tephritidae: Diptera) complex species collected in and around the Ryukyu islands of Japan using the mitochondrial A-T rich control region, Results Bull. Plant Prot. Jpn. 46 (2010) 17−23.

[53] A. Lara, J.L. Ponce de León, R. Rodríguez, D. Casane, G. Côté, L. Bernatchez, et al., DNA barcoding of Cuban freshwater fishes: evidence for cryptic species and taxonomic conflicts, Mol. Ecol. Resour. 10 (2010) 421−430, http://dx.doi.org/10.1111/j.1755-0998.2009.02785.x.

[54] K.A. Shufran, J.D. Burd, J.A. Anstead, G. Lushai, Mitochondrial DNA sequence divergence among greenbug (Homoptera: aphididae) biotypes: evidence for host-adapted races, Insect Mol. Biol. 9 (2000) 179−184, http://dx.doi.org/10.1046/j.1365-2583.2000.00177.x.

[55] F.C. Kafatos, A. Efstratiadis, B.G. Forget, S.M. Weissman, Molecular evolution of human and rabbit beta-globin mRNAs, Proc. Natl. Acad. Sci. U. S. A. 74 (1977) 5618−5622.

[56] T. Miyata, T. Yasunaga, T. Nishida, Nucleotide sequence divergence and functional constraint in mRNA evolution, Proc. Natl. Acad. Sci. U. S. A. 77 (1980) 7328−7332, http://dx.doi.org/10.1073/pnas.77.12.7328.

[57] K.P. Johnson, J. Seger, Elevated rates of nonsynonymous substitution in island birds, Mol. Biol. Evol. 18 (2001) 874−881, http://dx.doi.org/10.1093/oxfordjournals.molbev.a003869.

[58] J.J. Wernegreen, N.A. Moran, Evidence for genetic drift in endosymbionts (Buchnera): analyses of protein-coding genes, Mol. Biol. Evol. 16 (1999) 83−97.

[59] W.M. Brown, The mitochondrial genome of animals, Mol. Evol. Genet. (1985) 95−130, http://dx.doi.org/10.1038/mt.2008.22. R.J. Macintyre.

[60] J.C. Fay, C.-I. Wu, Sequence divergence, functional constraint, and selection in protein evolution, 2003, http://dx.doi.org/10.1146/annurev.genom.4.020303.162528.

[61] W. Delport, K. Scheffler, C. Seoighe, Models of coding sequence evolution, Br. Bioinform 10 (2009) 97−109, http://dx.doi.org/10.1093/bib/bbn049.

[62] P. Gomez-Polo, O. Alomar, C. Castañé, J. Riudavets, N. Agustí, Identification of Orius spp. (Hemiptera: Anthocoridae) in vegetable crops using molecular techniques, Biol. Control 67 (2013) 440−445, http://dx.doi.org/10.1016/j.biocontrol.2013.09.017.

[63] J.F. Germain, C. Chatot, I. Meusnier, E. Artige, J.Y. Rasplus, a Cruaud, Molecular identification of Epitrix potato flea beetles (Coleoptera: Chrysomelidae) in Europe and North America, Bull. Entomol. Res. 103 (2013) 354−362, http://dx.doi.org/10.1017/S000748531200079X.

[64] S. Kumar, A.J. Filipski, F.U. Battistuzzi, S.L. Kosakovsky Pond, K. Tamura, Statistics and truth in phylogenomics, Mol. Biol. Evol. 29 (2012) 457−472, http://dx.doi.org/10.1093/molbev/msr202.

[65] C. Vadivalagan, P. Karthika, K. Murugan, C. Panneerselvam, M. Paulpandi, P. Madhiyazhagan, H. Wei, A.T. Aziz, M.S. Alsalhi, S. Devanesan, M. Nicoletti, R. Paramasivan, D. Dinesh, G. Benelli, Genetic deviation in geographically close populations of the dengue vector *Aedes aegypti* (Diptera: Culicidae): influence of environmental barriers in South India, Parasitol. Res. 115 (2016) 1149−1160.